

**Japanese Laid-open Patent Application Publication No. Hei-8-293039 (JP-8-293039A)**

(54) Title of the Invention: Musical Image Conversion Device

(51) Int Cl<sup>6</sup> : G06T 13/00, 11/80

G06F 15/62

G09F 27/00

G09F 27/00

G10H 1/00, 1/38

G10H 1/00, 1/38

G06F 15/62

(43) Publication Date: November 5, 1996

(21) Application No.: Hei-7-98208

(22) Application Date: April 24, 1995

(72) Inventor: SATO, Junichi

c/o MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.

No. 1006, Kadoma, Kadoma City, Osaka Pref.

(71) Applicant: 000005821

MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.

(74) Attorney: TAKIMOTO, Tomoyuki, et al.

**Abstract**

[Purpose] To convert music into images or visual art having reality.

[Constitution] In the device having a motion database 107 for storing the motion of an articulated object, a motion generating section 106, on the basis of the chord detected by a chord detecting section 105, searches motion of the articulated object from the motion database 107 and generates motion data when the volume buildup point is detected by a peak detecting section 103.

[Effect] it is possible to generate, from music, images having definite shape and motion, such as the motion of a human dancing to the music.

***Page2, left column, line 36 to the end of the specification***

**Specification**

[Detailed Description of the Invention]

[0001]

[Field of the Invention]

The present invention relates to a musical image conversion device for converting musical information into images.

[0002]

[Prior Art]

As an example of a prior art of converting sounds into images, a technique is disclosed in which the motion of sounds in each part of a music is made to correspond to the motion of images having a basic shape (JP-A-63-184875).

[0003]

[Problems to be Solved by the Invention]

However, the prior art above has a drawback. That is, since the motions of sounds are made to correspond directly to the motions of images, the images obtained are abstract without relevant motions as a whole, and the visual art intended by the related art ends up in being abstract art.

[0004]

The object of the present invention is to provide a musical image conversion device for converting music or acoustic art, into images or visual art having reality.

[0005]

[Means for Solving the Problem]

To achieve the above object, the present invention comprises: a motion database for storing the motion of an articulated object, a peak detecting section for generating timing to switch the motion from musical signals, a chord detecting section for detecting chords from musical signals, and a motion generating section for selecting motions from the motion database based on the chords in time to the timing to switch the motion.

[0006]

[Function]

A certain number of motions made within a certain period of time by an articulated object such as a human are stored in advance in a database. With a sound buildup point as a cue, a motion is searched according to the chord constitution of that part. Musical information is converted into a series of pieces of motion information by appropriately switching the motion at every point of sound buildup. Motions of the articulated object are applied to the image of the articulated object to display an animated image.

[0007]

[Embodiment]

(First embodiment)

An embodiment of the present invention will be described below in reference to appended drawings. FIG. 1 shows the constitution of an embodiment of the present invention.

[0008]

First, music is inputted as electric signals of voice through a musical signal inputting section 101. A volume detecting section 102 takes sound volume information out of musical information inputted through the musical signal inputting section 101. In the case the musical information inputted is electric signals of voice, since the square of that value is the volume information, its envelope line can be obtained as the volume information signal. A peak detecting section 103 detects a volume buildup point from the volume information signal obtained with the volume detecting section 102. Arranging such that only the buildup exceeds an average volume for a certain period of time (about 2 or 3 seconds) immediately before, detection of small sound buildups can be prevented.

[0009]

A tone interval detecting section 104 detects tone interval information from the musical information inputted through the musical signal inputting section 101. Electric signals of voice are subjected to frequency analysis using Fourier transform. Tone intervals used in music are the values

expressed with the equation 1:

[0010]

(Equation 1)

$$f(n, m) = 440 \times 2^{n/12} \times 2^m \text{ (Hz)}$$

where,

$n$ : 0, ..., 11

$m$ : integer

[0011]

The basic tone  $I_a$  of 440 Hz is  $f(0, 0)$ . The letters  $n$  and  $m$  denote the intervals by a half tone and an octave, respectively. Therefore, the frequency component  $I(n, m)$  of the frequency  $f(n, m)$  expressed with the equation 1 is obtained with the tone interval detecting section 104. A chord detecting section 105 detects a chord according to the tone interval information obtained with the tone interval detecting section 104 at the volume buildup point obtained with the peak detecting section 103. First, a tone that is displaced by exactly an octave is deemed to be the same tone and their components  $I(n)$  are calculated with the equation 2:

[0012]

(Equation 2)

$$I(n) = \sum_m I(n, m)$$

[0013]

A certain number of  $n$ 's are taken out in the order of decreasing magnitude of component  $I(n)$  and combinations of the  $n$ 's are detected as chords.

[0014]

A motion generating section 106, on the basis of the information on the volume buildup point detected with the peak detecting section 103 and the chord detected with the chord detecting section 105, searches motion of the articulated object from the motion database 107, and synthesizes. FIG. 4 shows an example in which a human is deemed to be an articulated object. Motions of limbs of the articulated object shown in FIG. 4 are represented with a series of numerical values of position coordinates and angles of joints, and stored in advance in the motion database 107. A volume buildup point detected with the peak detecting section 103 is made a motion changing point. The kinds of chords detected with the chord detecting section 105 are matched in advance with the motion data stored in the motion database 107, and the motion occurring after the motion changing point is searched by the kind of chords. Joint position coordinates or joint angle value series stored as motion data are made to change smoothly from the precedent motion by applying a spline processing or the like. A motion output section 108 visualizes to display the motion data generated with the motion generating section 106.

[0015]

As a result, an articulated object moving in time to the music may be displayed using the pre-stored motion data.

[0016]

(Second embodiment)

Next, another embodiment of the motion generating section 106 according to the present invention is described in reference to FIG. 2. A chord classifying section 201 classifies into kinds the chords detected with the chord detecting section 105.

[0017]

An example of determining the kinds of chords with the chord classifying section 201 is described in reference to the flowchart of FIG. 5. Here, three n's (m1, m2, and m3) are selected with which chord components are great, assuming  $m1 < m2 < m3$ . In the case the values of (m2-m1) and (m3-m2) prove to be 4 and 3 respectively as a result of checking the values, the chord is of the major key. In the case they prove to be 3 and 4, the chord is of the minor key. For example, in the case a chord of do, mi and sol is detected, since m1, m2, and m3 are 3, 7, and 10, respectively, the chord is determined to be the major key. In the case of la, do, and mi, since m1, m2, and m3 are 0, 3 and 7, the chord is determined to be the minor key. Some other determinations may be made by the two differential values above.

[0018]

In the case the determination cannot be made, the chord is rotated by replacing the m2, m3 and n1 + 12 used so far with new n1, m2, and m3, and determination is made. For example, if a chord of la, do and fa is present, the n1, m2 and m3 are 0, 3 and 8, respectively, so it cannot be determined to be either major or minor key. So determination is made by rotating the chord with new n1, m2 and m3 of 3, 8 and 12, respectively. If the determination cannot be made, one more rotation is made and determined. If the determination fails three times, it is determined to be 'unidentified' and the process is finished.

[0019]

In FIG. 2, a motion searching section 202 searches motion data corresponding to the chord classification results obtained with the chord classifying section 201 from the motion database 107. A motion synthesizing section 203 smoothes by the spline processing or the like the motion data searched with the motion searching section 202 and the motion data found so far and synthesizes so that the motions are consistent. As a result, the motion of the articulated object may be varied according to the atmosphere, the contents of the musical chord.

[0020]

(Third embodiment)

Next, still another embodiment of the motion generating section 106 according to the present invention is described in reference to FIG. 3. A bar recognizing section 301 determines a motion change point based on the volume buildup point detected with the peak detecting section 103.

[0021]

An operation sample of the bar recognizing section 301 is described in reference to FIG. 6. A bar is defined as the shortest unit time for performing one motion. The graph shows a volume buildup time point detected with the peak detecting section and the volume at that time point. The volume is indicated along the vertical axis of the graph while time is along the horizontal axis. It is assumed that the inputted musical information represents a rhythm that is distinct to some extent and the volume buildup point is approximately one of time points of 0, 1a, 2a, ... In the drawing, 'bar time' refers to a period of time for performing one motion, 'elapsed time within a bar' refers to the time

elapsed within a bar, and 'bar head time' refers to the time point at the head of a bar.

[0022]

It is assumed that the volume first builds up at the time point 0. A motion starts from that point. The next buildup time point is assumed to be 'a.' When the volume at the time point 'a' is greater than that at the time point 0, the time point 'a' is defined as the reference time point; otherwise the time point 0 is defined as the reference time point. And the period of time 'a' is defined as the unit time. As the unit time is determined at the time point 'a,' the bar time at that time point is 'a.' When the time point 'a' is assumed to be the reference time point, the time elapsed within the bar is 0, and the bar head time is 'a.'

[0023]

At the time point '2a,' there is no input and so it is deemed to be in the middle of the bar. Therefore, the bar time is 2a, the time elapsed within the bar is 'a,' and the bar head time remains unchanged, 'a.'

[0024]

At the time point '3a,' although there is an input, since the input is rather small, it is hard to determine whether it denotes the head or intermediate point of the bar. Therefore, that time point is determined to be 'at the head or in the middle' of the bar. The bar time is determined to be 2a or 3a, and the time elapsed within the bar is determined to be 0 (the bar has changed) or 2a (the bar has not changed). Since input is small also at the time point 4a, the state at the time point 3a is carried over.

[0025]

When a great input is given at the time point 5a, that point may be deemed to be the head of a bar, and the inputs at the time points 3a and 4a are determined to have been given in the middle of the bar. Therefore, the bar time may be deemed to be 4a, and the time elapsed within the bar is deemed to be 5a.

[0026]

Thereafter, in the case a great input is given every time a period of 4a elapses, it is deemed there is no contradiction, and the bar time 4a is maintained. The sound volume is judged according to a reference which is the average of the sound volume during a preceding period of a few seconds. The bar recognizing section 301 outputs the bar head time. The unit time 'a' is dynamically set according to the volume buildup points of several preceding bars.

[0027]

The motion synthesizing section 203, according to the time span between motion change points determined with the bar recognizing section 301, determines at what speed the motions are to be synthesized. For example, in case a motion reproduction time T is 4a, motion data must be interpolated and corrected so that one motion is finished over the time T=4a. When it is assumed that the motion data stored in the motion database 107 are for M frames per one motion and that the display system requires motion data of D frames per second, in order to display the d-th frame on the display system, the motion data of the m-th frame should be used as shown in (Equation 3).

[0028]

(Equation 3)

$$m = \frac{Md}{TD}$$

[0029]

Since  $m$  is not an integer generally, motion data for display are obtained by interpolating a motion data of (integer portion of  $m$ )-th frame and a motion data of (integer part of  $m$  plus 1)-th frame.

[0030]

As a result, it is possible to roughly set the motion change point according to the speed of the piece of music and to generate motions according to the music.

[0031]

(Fourth embodiment)

Next, an embodiment of the motion output section 108 according to the invention is described. The motion output section 108 displays the motion data outputted with the motion generating section 106 by applying them to the three-dimensional shape model of an articulated object generated by computer graphics. A three-dimensional shape model of a human shape as shown in FIG. 4 for example is used as the articulated object, and joint position coordinates and joint angles of the motion data are applied to the three-dimensional shape model.

[0032]

As a result, the manner of motion, in time to the music, of an articulated object such as a human having a definite shape may be visualized.

[0033]

(Fifth embodiment)

Next, the fifth embodiment of the invention is described. As music information, signals corresponding to performance control information, such as MIDI control signals are inputted to the musical signal inputting section 101. The MIDI control signals have interval information and sound volume information for individual tones. The volume detecting section 102 may take out volume information as a piece of music by calculating the sum of sound volume information of tones present at timing points for which tone producing information is present. Since it is easy to take out interval information with the tone interval detecting section 104, thereafter it is possible to carry out a similar process to that when electric signals of voice are inputted as musical information.

[0034]

Thus, load of interval detection and volume detection processes is alleviated.

[0035]

[Effect of the Invention]

According to the invention as described above, it is possible to convert the acoustic art, the music, into specific visual art, the motion of articulated object having a definite shape, for example the motion of a human dancing to the music.

[Brief Description of Drawings]

FIG. 1 is the block diagram of the constitution of an embodiment of the present invention.

FIG. 2 is the block diagram of an example constitution of the motion generating section of the embodiment.

FIG. 3 is the block diagram of another example constitution of the motion generating section of the embodiment.

FIG. 4 is a line drawing of a human as an example of articulated object of the embodiment.

FIG. 5 is a flowchart of the chord classification procedure of the embodiment.

FIG. 6 shows the function of the bar recognizing section which determines motion change points using volume data in time sequence.

**[Description of Reference Numerals]**

**101: Musical signal input section**

**102: Volume detecting section**

**103: Peak detecting section**

**104: Tone interval detecting section**

**105: Chord detecting section**

**106: Motion generating section**

**107: Motion database**

**108: Motion output section**

**201: Chord classifying section**

**202: Motion detecting section**

**203: Motion synthesizing section**

**301: Bar recognizing section**

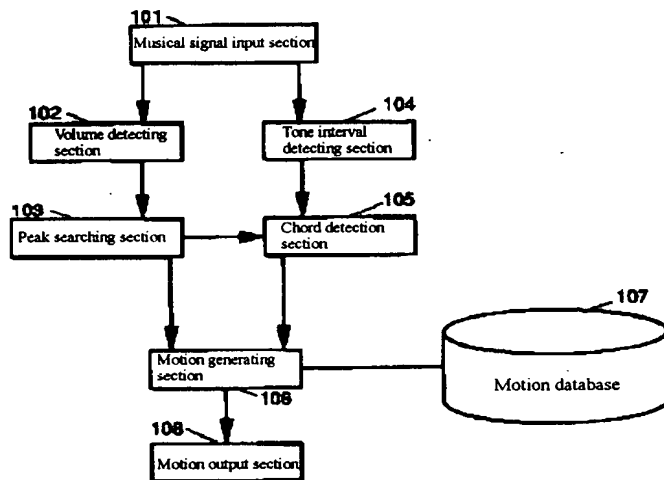


FIG. 1

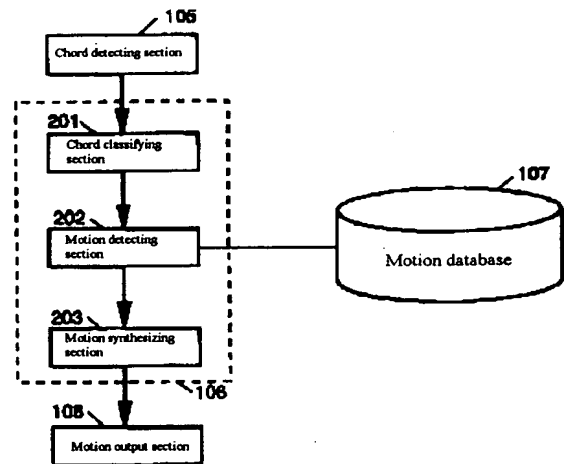


FIG. 2

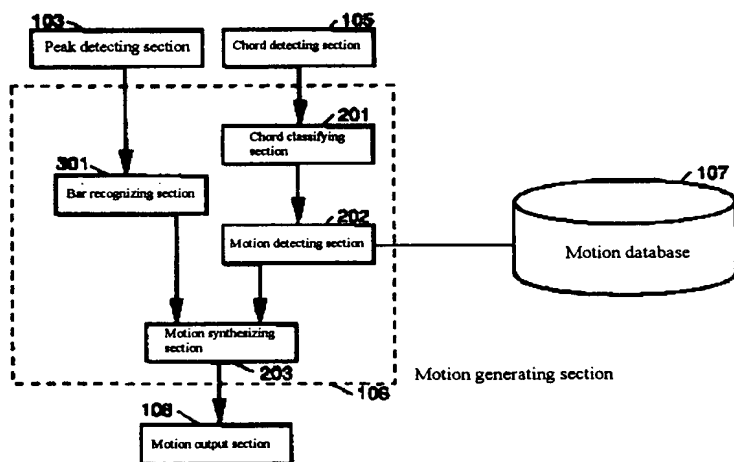


FIG. 3

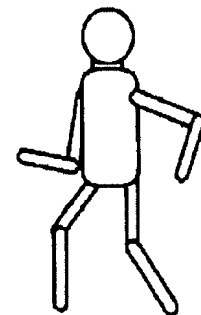


FIG. 4



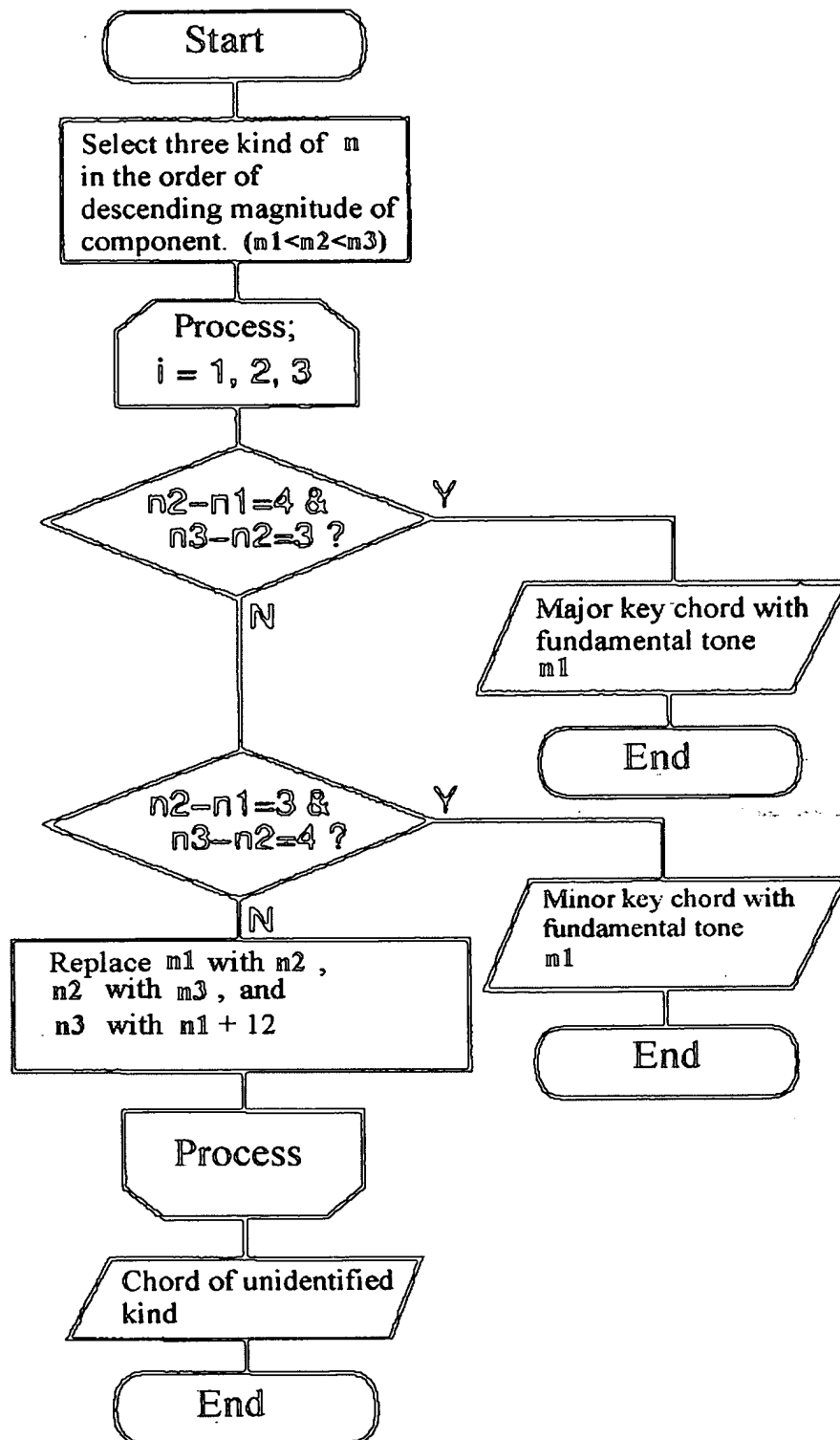


FIG. 5

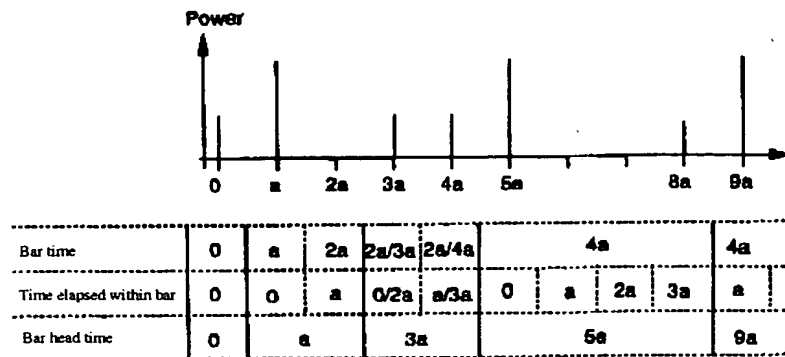


FIG. 6

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-293039

(43) 公開日 平成8年(1996)11月5日

(51) Int.Cl. <sup>8</sup>	識別記号	庁内整理番号	FI	技術表示箇所
G 0 6 T 13/00			G 0 6 F 15/62	3 4 0 D
11/80			G 0 9 F 27/00	N
G 0 9 F 27/00			G 1 0 H 1/00	Z
G 1 0 H 1/00			1/38	Z
1/38			G 0 6 F 15/62	3 2 1 A
			審査請求 未請求 請求項の数5 OL (全6頁)	

(21) 出願番号 特願平7-98208

(22) 出願日 平成7年(1995)4月24日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 佐藤 潤一

大阪府門真市大字門真1006番地 松下電器

産業株式会社内

(74) 代理人 弁理士 滝本 智之 (外1名)

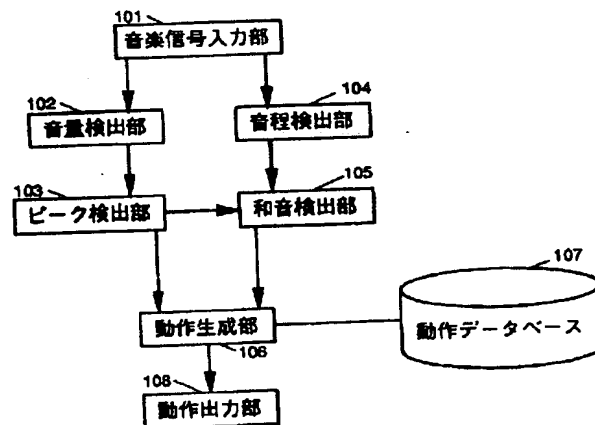
(54) 【発明の名称】 音楽画像変換装置

(57) 【要約】

【目的】 音楽を、具体性を持った視覚芸術としての画像に変換する。

【構成】 関節物体の動作を記憶する動作データベース107を有し、動作生成部106は、ピーク検出部103で音量の立ち上がり箇所が検出された時点に、和音検出部105で検出された和音に基づいて動作データベース107から動作を検索し、動作データを生成する。

【効果】 音楽に合わせて踊る人間の画像のような、具体的な形状や動きを持つ画像を、音楽をもとに生成することができる。



## 【特許請求の範囲】

【請求項 1】音楽情報を入力する音楽信号入力部と、前記音楽信号入力部で入力された音楽情報から音高を検出する音高検出部と、前記音高検出部で検出された音高データから音の立ち上がり箇所を検出するピーク検出部と、前記音楽信号入力部で入力された音楽情報から音程の分布を検出する音程検出部と、前記ピーク検出部で検出された音の立ち上がり箇所の時点において、前記音程検出部で検出された音程分布から和音を検出する和音検出部と、関節物体の動作を記憶する動作データベースを保持し、前記ピーク検出部で検出された音の立ち上がり箇所のタイミングと前記和音検出部で検出された和音とをもとに動作データベースから動作データを取り出して動作を生成する動作生成部と、前記動作生成部で生成された動作を出力する動作出力部とから構成されることを特徴とする音楽画像変換装置。

【請求項 2】動作生成部は、和音検出部で検出された和音データを分類する和音分類部と、前記和音分類部で分類された結果をもとに、前記動作データベースから動作データを取り出す動作検索部と、前記動作検索部で取り出された複数の動作データを接合して動作を合成する動作合成部とから構成されることを特徴とする音楽画像変換装置。

【請求項 3】動作生成部は、ピーク検出部で検出された音の立ち上がり箇所のタイミング列から小節単位を認識する小節認識部を有し、前記小節認識部で決定された小節単位の時間に行う動作を動作検索部で検索し、動作合成部は、その動作データ列から動作を合成することを特徴とする請求項 2 記載の音楽画像変換装置。

【請求項 4】動作出力部は、動作生成部で生成された動作を行う関節物体の画像を出力することを特徴とする請求項 1 記載の音楽画像変換装置。

【請求項 5】音楽信号入力部は、音楽の演奏制御情報に相当する信号を入力することを特徴とする請求項 1 記載の音楽画像変換装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は音楽情報を画像に変換する音楽画像変換装置に関するものである。

## 【0002】

【従来の技術】音を画像に変換する従来技術の例としては、曲の各パートの音の動きを基本形状を持つ画像の動きに対応させる手法（特開昭 63-184875 号公報）があった。

## 【0003】

【発明が解決しようとする課題】しかしながら、上記従来の技術では音の動きを画像の動きに直接対応させるため、得られる画像は、画像全体として関連性のない動きを持った抽象的なものとなり、上記従来技術が目的とする視覚芸術が、抽象芸術に限られてしまうという欠点を

有していた。

【0004】本発明の目的は、聴覚芸術としての音楽を、具体性を持った視覚芸術としての画像に変換する音楽画像変換装置を提供することにある。

## 【0005】

【課題を解決するための手段】この目的を達成するため本発明は、関節物体の動作を記憶する動作データベースを有し、音楽信号から動作を切り替えるタイミングを生成するピーク検出部と、音楽信号から和音を検出する和音検出部と、動作を切り替えるタイミングで和音をもとに動作データベースから動作を選択する動作生成部とを有する。

## 【0006】

【作用】人間などの形状を有する関節物体がある時間内に行う動作を、あらかじめいくつか動作データベースに保存しておき、音の立ち上がり箇所をきっかけに、その部分の和音構成に基づいて動作を検索する。音の立ち上がり箇所ごとに動作を適当に切り替えることにより、音楽情報を一連の動作情報に変換し、関節物体の画像にその動作を適用して画像化する。

## 【0007】

## 【実施例】

（実施例 1）以下本発明の実施例について、図面を参照しながら説明する。図 1 は本発明の一実施例における構成図である。

【0008】まず、音楽信号入力部 101 により音楽を音声の電気信号として入力する。音高検出部 102 は、音楽信号入力部 101 により入力される音楽情報から音高の情報を取り出す。入力された音楽情報が音声の電気信号の場合は、その 2 乗の値が音高情報となるので、その包絡線を音高情報信号として得る。ピーク検出部 103 は、音高検出部 102 で得られる音高情報信号から、音高の立ち上がり箇所を検出する。直前のある時間

（2、3 秒程度）での平均音高を超える立ち上がりのみを検出するようにすることにより、細かい音の立ち上がりを検出するのを防ぐことができる。

【0009】音程検出部 104 は、音楽信号入力部 101 により入力される音楽情報から、音程の情報を検出する。音声の電気信号をフーリエ変換を用いて周波数解析を行う。音楽で用いられる音程は通常（数 1）で表せる値となる。

## 【0010】

## 【数 1】

$$f(n, m) = 440 \times 2^{n/12} \times 2^m \quad [\text{Hz}]$$

$$n: 0, \dots, 11$$

$$m: \text{整数}$$

【0011】 $f(0, 0)$  が 440 Hz の基準となるラの音で、 $n$  は半音間隔の音程を、 $m$  はオクターブの音程

を表す。したがって、(徴1)で表される周波数 $f$  ( $n, m$ )の周波数成分 $I(n, m)$ を音程検出部104で得る。和音検出部105は、ピーク検出部103で得られた音 $\square$ の立ち上がり箇所において、音程検出部104で得られた音程情報に基づいて和音の検出を行う。まず、ちょうどオクターブずれた音は同じ音とみなし、それらの成分 $I(n)$ を(徴2)のように計算する。

【0012】

【徴2】

$$I(n) = \sum_m I(n, m)$$

【0013】成分 $I(n)$ の大きい順にある個数の $n$ を取り出し、その $n$ の組み合わせを和音として検出する。

【0014】動作生成部106は、ピーク検出部103で検出された音 $\square$ の立ち上がり箇所の情報と、和音検出部105で検出された和音をもとに、動作データベース107から関節物体の動作を検索し、合成する。図4に、人間を関節物体とみなした例を示す。図4のような関節物体に対して、手足を動かしたときの動きを、関節の位置座標、あるいは関節の角度の値列として動作データベース107に記憶しておく。ピーク検出部103で検出された音 $\square$ の立ち上がり箇所を動作の変化点とする。和音検出部105で検出された和音の種類と動作データベース107の記憶する動作データとを対応づけておき、動作の変化点の時点以降の動作を和音の種類で検索する。動作データとして持っている関節位置座標あるいは関節角の値列は、スプライン処理などを行って変化前の動作から平滑に変化するようにする。動作出力部108は動作生成部106の生成した動作データを可視化して表示する。

【0015】以上の結果、あらかじめ記憶してある動作データをもとに、音楽に合わせて動く関節物体を表示することができる。

【0016】(実施例2)次に本発明における動作生成部106の実施例について、図2を参照しながら説明する。和音分類部201は、和音検出部105が検出した和音の種類を分類する。

【0017】和音分類部201における和音の種類判別の動作例を、図5のフローチャートを参照しながら説明する。ここでは、和音の成分 $I(n)$ が大きい $n$ を3つ選択し、 $n1, n2, n3$  ( $n1 < n2 < n3$ )とする。 $n2 - n1$ と $n3 - n2$ の値を調べ、それぞれ4、3であればそれは長調の和音であると判別でき、それぞれ3、4であれば単調の和音であると判別できる。例えば、ドミソの和音が検出された場合、 $n1, n2, n3$ はそれぞれ3、7、10であるので、長調の和音と判別でき、ラドミの和音の場合は $n1, n2, n3$ はそれぞれ0、3、7であるので単調の和音であると判別できる。その他にもこの2つの差の値からいくつかの判別が可能である。

【0018】もし判別ができなかった場合、新たな $n1, n2, n3$ を、それまで使っていた $n2, n3, n1 + 12$ にそれぞれ置き換えて和音を巡回し、判別を行う。例えば、ド、ファという和音があった場合は、 $n1, n2, n3$ はそれぞれ0、3、8となるので長調とも単調とも判別できないので、新たに $n1, n2, n3$ をそれぞれ3、8、12と巡回して判別を行う。これで判別ができない場合はもう一度巡回、判別を行い、3度目で判別ができない場合は「不明」と判別し、処理を終了する。

【0019】図2において、動作検索部202は、和音分類部201で得られた和音の分類の結果からそれに対応する動作データを動作データベース107から検索する。動作合成部203は動作検索部202が検索した動作データとそれまでの動作データとを、スプライン処理などで平滑化を行い、一連の動作となるように合成する。その結果、関節物体の動作を、音楽の和音の内容、つまり音の雰囲気によって変更することが可能となる。

【0020】(実施例3)次に本発明における動作生成部106の他の実施例について、図3を参照しながら説明する。小節認識部301は、ピーク検出部103が検出した音 $\square$ の立ち上がり箇所をもとに、動作変化点を決定する。

【0021】小節認識部301の動作例を、図6を参照しながら説明する。小節を、1動作を実行させる最小単位の時間と定義する。グラフは、ピーク検出部が検出した音 $\square$ の立ち上がり箇所の時刻とそのときの音 $\square$ の大きさを表す。グラフの縦軸は音 $\square$ を、横軸は時間を表す。入力されたのがある程度リズムのはっきりした音楽情報であり、近似的に音 $\square$ の立ち上がり箇所が時刻0、1a、2a……のいずれかであるものと仮定する。また図中、「小節時間」は1動作を実行させる時間を、「小節内経過時間」は小節内で経過した時間を、「小節先頭時刻」はその小節の先頭時刻を表す。

【0022】最初に音 $\square$ の立ち上がるのが時刻0であるとする。ここから動作がスタートする。次に立ち上がる時刻をaとする。時刻aの音 $\square$ が時刻0よりも大きい場合は時刻aを、そうでない場合は時刻0を基準時刻と定義する。また時間aを単位時間と定義する。時刻aの時点で単位時間が決定したため、その時点の小節時間はaとなる。また、時刻aを基準時刻とした場合は小節内経過時間は0、小節先頭時刻はaとなる。

【0023】時刻2aでは、入力がないため、そこは小節の途中であると判断される。したがって小節時間は2aとなり、小節内経過時間はaとなり、小節先頭時刻は変わらずaとなる。

【0024】時刻3aでは入力があるが、やや小さい入力のため、それが小節の先頭を表すのか小節の途中を表すのか判断しにくい。そこで、その時点は小節の「先頭か途中」であると判断する。小節時間は2aか3aと

し、小節内経過時間は0（小節が変わった）か2a（小節が変わっていない）とする。時刻4aでも入力が小さいため、時刻3aの状態を継承する。

【0025】時刻5aで大きい入力が入ると、そこを小節の先頭とみることができ、時刻3a、4aの入力は小節の途中であると判断できる。そこで小節時間は4a、小節内先頭時刻を5aとすることができる。

【0026】以降、時刻が4aを経過するごとに大きい入力があれば、矛盾はないとして小節時間4aが維持される。音圧の大きさは、それ以前の微秒間の音圧の平均値を基準にして判断する。小節認識部301は、小節先頭時刻を出力する。また単位時間aは、各時刻以前の微小節間の音圧立ち上がり箇所をもとに、動的に決定される。

【0027】動作合成部203は、小節認識部301が決定した動作変化点どうしの時間をもとに、動作をどれくらいの速度で合成すればよいか決定する。例えば動作再生時間Tが4aであれば、 $T=4a$ の時間で一動作を終了するように、動作データを補間、修正する必要がある。動作データ107に記憶された動作データが1動作当たりMコマ分のデータを持ち、表示系が1秒あたりDコマの動作データを必要とするとすれば、表示系でdコマ目の表示を行うためには（数3）の通りmコマ目の動作データを用いればよい。

【0028】

【数3】

$$m = \frac{Md}{TD}$$

【0029】一般にはmは整数とならないので、（mの整数部）コマ目と（mの整数部）+1コマ目の動作データの補間により、表示する動作データを得る。

【0030】この結果、動作の変化点を曲のスピードに合わせて大まかにとることができ、曲に合わせて動作を生成することが可能となる。

【0031】（実施例4）次に本発明における動作出力部108の実施例について説明する。動作出力部108は動作生成部106が出力した動作データを、コンピュータグラフィックスで生成する関節物体の3次元形状モデルに適用することで表示する。関節物体として、例えば図4のように人間の形状を表す3次元形状モデルを用い、動作データが持つ関節位置座標や関節角度をこの3次元形状モデルに適用していく。

【0032】この結果、人間など具体的な形状を持つ関節物体が、音楽に合わせて動くさまを映像化することが

可能となる。

【0033】（実施例5）次に本発明の実施例5について説明する。音楽情報として、音楽の演奏制御情報に相当する信号、例えばMIDI制御信号を音楽信号入力部101に入力する。MIDI制御信号は音程情報、音圧情報を個々の音について持っている。音圧検出部102では、発音情報が存在するタイミングにおいて、そのタイミングに存在する音の音圧情報の総和を計算することで、音楽としての音圧情報を取り出すことができる。音程検出部104で音程情報を取り出すことは容易であるので、それ以降は音楽情報として音声の電気信号を入力した場合と同様の処理を行うことができる。

【0034】以上の結果、音程検出や音圧検出の処理の負荷を軽減することができる。

【0035】

【発明の効果】以上のように本発明によれば、音楽という聴覚芸術を、人間が音楽に合わせて踊るような、具体的な形状を持った物体の動きという、具体性を持った視覚芸術に変換することが可能となる。

【図面の簡単な説明】

【図1】本発明の一実施例の構成を示すブロック図

【図2】本実施例の動作生成部の構成例を示すブロック図

【図3】本実施例の動作生成部の他の構成例を示すブロック図

【図4】本実施例の人間を関節物体の一例として示す線図

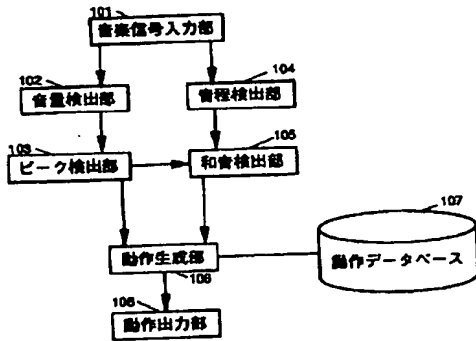
【図5】本実施例の和音分類部の手続きを示すフローチャート

【図6】時系列での音圧データから動作の変化点を決定する小節認識部の動作を示す図

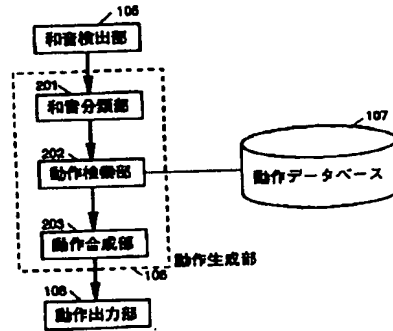
【符号の説明】

- 101 音楽信号入力部
- 102 音圧検出部
- 103 ピーク検出部
- 104 音程検出部
- 105 和音検出部
- 106 動作生成部
- 107 動作データベース
- 108 動作出力部
- 201 和音分類部
- 202 動作検索部
- 203 動作合成部
- 301 小節認識部

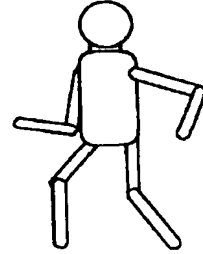
【図 1】



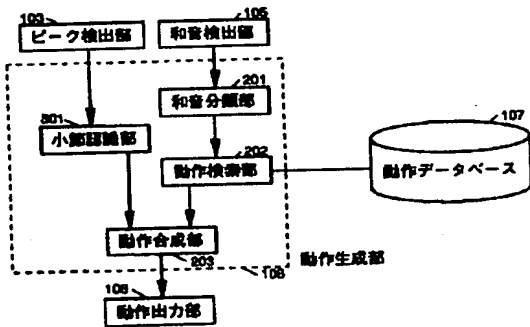
【図 2】



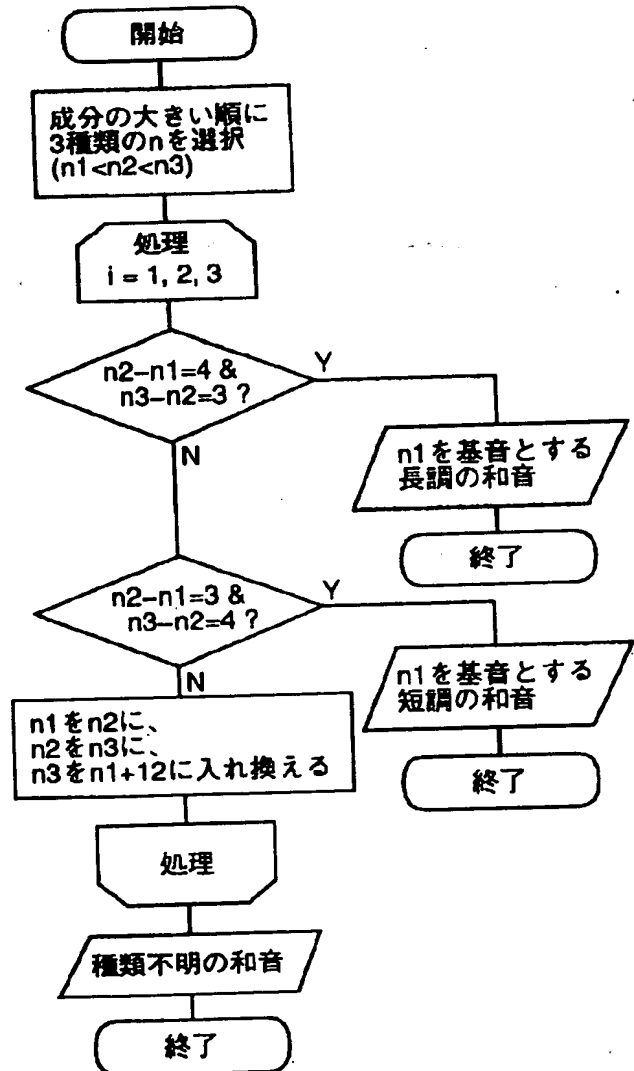
【図 4】



【図 3】



【図 5】



(6)

特開平8-293039

【図6】

